

# Gerald Amiel Ballena

✉ gmballena@up.edu.ph     github.com/GABallena     linkedin.com/in/gerald-amiel-ballena  
📍 Metro Manila, Philippines     orcid.org/0009-0000-8857-9755

## Professional Summary

---

Software Engineer and Bioinformatics Specialist with extensive experience developing computational pipelines for sequencing data and designing reproducible workflows. Proficient in Python, data visualization, and high-throughput data analysis, with a demonstrated ability to work on interdisciplinary bioinformatics projects, including 16S RNA sequencing.

## Experience

---

### 2024 – Present

### Project Technical Specialist

*University of the Philippines, College of Public Health*

- Developed scalable pipelines for high-throughput sequencing data (2.5 TB each), including preprocessing, assembly, and taxonomic profiling using **Snakemake**.
- Designed Python scripts for analyzing sequencing data, automating tasks, and generating insightful visualizations.
- Collaborated with interdisciplinary teams, contributing to public health surveillance and environmental microbiology projects.
- Enhanced reproducibility by implementing automated documentation and detailed logging systems.
- Applied statistical analyses to sequencing data using **pandas** and **scikit-bio**.

### 2022 – 2024

### Graduate Researcher

*University of the Philippines Diliman*

- Built bioinformatics workflows for metagenomic analysis, including k-mer profiling, 16S RNA analysis, and antimicrobial resistance (AMR) gene prediction.
- Implemented automated diversity profiling using **Python** and **R** for public health studies.
- Developed robust pipelines for taxonomic profiling (**Kraken2**, **Bracken**) and genome assembly (**MEGAHIT**, **metaSPAdes**).
- Conducted exploratory data analyses and visualized results using **matplotlib** and **ggplot2**.

## Education

---

### University of the Philippines Diliman.....

**Graduated:** July 2022

**Thesis:** In silico assessment of the association of pathogenicity and metal-resistance potential of *Fusarium* spp. .... Pre-print Link

**GPA:** 1.72 (DOST ASTHRDP Scholar)

### University of the Philippines Baguio.....

**Graduated:** June 2018

**Thesis:** Bioelectrocatalysis by Novel Electrogenic Alkaliphilic Bacteria *Bacillus* sp. BAB-3442 Using Dual-Chambered Microbial Fuel Cell

**Achievements:** Advanced Placement in Advanced Algebra

## Technical Skills

---

### Programming Languages.....

Python (Advanced), R (Intermediate), Bash (Intermediate), Perl (Intermediate)

### Bioinformatics Expertise.....

Taxonomic profiling (**Kraken2**, **Bracken**), genome assembly (**metaSPAdes**, **MEGAHIT**), k-mer profiling (**Jellyfish**, **MASH**), antimicrobial resistance gene detection (**RGI**)

### Workflow Automation.....

Snakemake, Conda, Docker, Slurm

### Data Visualization and Analysis.....

**pandas**, **scikit-learn**, **matplotlib**, **ggplot2**, **QGIS**

## Key Projects

---

### 16S RNA Analysis Pipeline:

- Designed pipelines for 16S RNA sequencing data analysis, focusing on diversity profiling and taxonomic classification.
- Automated read preprocessing and quality filtering using **Trimmomatic** and **Cutadapt**.
- Applied **pandas** and **scikit-bio** for data manipulation, statistical testing, and entropy calculations.

### Bioinformatics Tool Management:

- Developed a Python scraper to automate dependency discovery and version updates for bioinformatics tools from Bioconda.
- Managed YAML-based environments for streamlined tool installation and reproducibility.

### Visualization Pipelines:

- Generated histograms, violin plots, and ridgeline plots to visualize diversity metrics using **ggplot2** and **matplotlib**.
- Created heatmaps for beta diversity metrics and non-metric multidimensional scaling (NMDS) plots.

## Certifications

---

**Data Science and Bioinformatics:** Completed advanced courses on machine learning workflows, Bioconductor, and differential expression analysis.

## Soft Skills

---

**Collaboration:** Proven ability to work effectively in cross-functional teams.

**Problem-Solving:** Adept at debugging and optimizing computational pipelines.

**Communication:** Skilled in presenting technical findings to both scientific and non-technical audiences.